REMARKS

Paragraph 3 of the Office Action rejected claims 1-5 and 14-19 as having been anticipated by Viikki. Claims 2 and 4 have been canceled. Of the remaining ones of these rejected claims, only claims 1 and 19 are independent. All the rest depend from Claim 1.

Claims 1 and 19 recite the training of acoustic phonemes by mapping utterances against phonetic spellings produced by a training pronunciation guesser. As amended, Claims 1 and 9 make clear that each of multiple training words is spoken by multiple speakers, and this is used to derive multiple speaker acoustic models for each of a set of phonemes, each of which multi-speaker acoustic models is derived from the sounds of utterances from multiple speakers mapped against a given phoneme in one or more of the phonetic spellings generated by the training pronunciation guesser. This is very distinct from what is recited in Viikki, which merely use a pronunciation guesser to generate a phonetic spelling for a new word to be used in the recognition of that word and in adaptation of the system's acoustic models by an individual speaker. There is no indication that the acoustic information used to generate Viikki's speaker-independent acoustic models are derived by mapping utterances of each word in a training set against a phonetic spelling of that word generated by a pronunciation guesser. There is no suggestion that the phonetic spellings used in the HMM training described in the portion of Viiki cited by the Office Action's rejection of claim 19 were generated by a pronunciation guesser.

Furthermore claims 1 and 19 have been amended to make it clear that not only is a training pronunciation guesser used to generate phonetic spellings use to train up multi-speaker acoustic models, but also that a recognition pronunciation guesser that makes somewhat similar phonetic spelling errors is used to create phonetic spellings used for recognition. It should be noted that is has generally been considered inappropriate to use pronunciation guessers to create the phonetic spellings used to

train up multi-speaker acoustic models, because the phonetic spelling errors created by such pronunciation guesser causes such acoustic models to be more accurate. But one of the major creative insight behind the invention recited in amended claims 1 and 19 is that the erroneous multi-speaker acoustic models created by training against phonetic spellings generated by a training pronunciation guesser, which normally decrease recognition performance, can actually improve recognition performance if the pronunciation guesser makes phonetic spelling errors similar to those made by a pronunciation guesser used to generated phonetic spellings used in recognition in conjunction with such multi-speaker acoustic models, including acoustic models that have been adapted from such multi-speaker acoustic models.

Therefore, it is respectfully requested that Claims 1 and 19, as amended, are clearly not anticipated by Viikki. Thus, it is respectfully requested that Claims 1 and 19, and all the remaining claims that depend from claim 1, that is claims 2, 5-10, and 12-18 which depend from claim 1 are patentable, and should be allowed.

Although claims 3, 5, and 14-18 that have been rejected as anticipated by Viikki are patentable because they depend from claim 1, it should be noted that these claims, like many of the other rejected dependent claims in this application, contain additional limitations that further argue for their allowance. The following are a few examples.

Claim 3 (as well as claim 19) recites "...5% or more of the occurrences of vowel phonemes placed in the phonetic spellings of the acoustic training words by the training pronunciation guesser are phonetic spelling errors...." There is no particular magic with exactly 5%, other than this is an error rate normally considered unacceptably high for the training of acoustic phoneme models, that is it is an error rate high enough that it would normally substantially decrease the usefulness of the acoustic models it is used to train, thus, teaching against the use of such an error-prone pronunciation guesser for generate phonetic spellings used for training. Contrary to what the Office Action says about Claim 3, it appears unlikely from Viikki that the recognition pronunciation guesser it discloses has a vowel spelling error rate as high as 5%. This is because Viikki's abstract says recognition errors due to both pronunciation guessing and automatic

language identification is approximately 4%, and the average decrease in recognition rate just due to language identification over the six languages shown in FIG. 2 is close to 3%, indicating that only about 1% of the recognition error is due to pronunciation guessing. In addition, as indicated above, with regard to claim 1, Viikki does not appear to show the use of phonetic spellings generated by pronunciation guessing in the training of multi-speaker acoustic models.

Amended Claim 17 recites that "...the set of training words are a representative distribution of names from US phone books." This means the set of training words for which the utterances from multiple speakers are mapped against phonetic spellings generated by the training pronunciation guesser is very large, and thus very different than the number of words that might be adapted by an individual user in a system such as Viikki's.

Paragraph 5 of the Office Action rejected Claims 6-10 and 20-25 over Viikki in view of the Sensory reference.

It is respectfully submitted that Claims 6-10 are patentable because they depend from amended Claim 1, and thus they should be allowable for the same reasons as Claim 1, which are stated above.

Similarly it is respectfully submitted that Claims 20-25 are patentable because they depend from amended Claim 19, and thus they should be allowable for the same reasons as Claim 19.

Paragraph 6 of the Office Action rejected claims 12-13, 26-39, and 47-60 over Viikki in view of Baker.

Claims 12-13 depend from Claim 1 and thus are patentable for the same reasons as is claim 1, stated above.

With regard to claim 26, and claims 27-30 that depend from it, as the Office action states, Baker does show phoneme-in-context acoustic models, but it does not appear to shown novel features recited in the last two paragraphs of claim 26. There are (1) having the acoustic models that represent a phoneme in a given phonetic context be represented by

> "a blended acoustic model that represents a given phoneme in a given phonetic context as a distribution of sounds corresponding to utterances of the given phoneme and utterances of an associated set of one or more other phonemes, where both the sounds corresponding to the utterances of the given phoneme and to utterances of one or more associated phonemes have each been derived from the utterances of multiple speakers..."

and (2) where

> "over the plurality of blended acoustic models, the relative weight allocated, in a given acoustic model representing a given phoneme in a given phonetic context, between sounds of utterances of the given phoneme and sounds of utterances of a specific one of the given phoneme's associated set of phonemes is correlated with the frequency with which the pronunciation guessing programming places the given phoneme in a position in a phonetic spelling in the given phonetic context where the correct phoneme for the position is said specific associated phoneme"

Note that roughly similar language is also included in independent Claims 31, 37. In all three of these independent claims this language has been modified from the original wording because the original wording was difficult to understand, and might actually not be grammatically correct. Also the wording has been amended to reflect that such blended models have been derived from multiple speakers.

Examples of such blended acoustic models and how they are weighted is described with regard to FIGS. 19-23 and 30 in the current application, and their associated text.

It should be appreciated that the acoustic blurring involved in the blended acoustic models recited in claims 26, 31, and 37 would normally be considered less accurate acoustic models, and thus conventional speech recognition approaches would teach away from their use. But as the inventors have realized, where the blurring of such models is made to correlate with the errors made by the pronunciation guesser that generates phonetic spelling used in recognition, such blurring can actually improve recognition performance significantly. The Office Action does not appear to make any discussion of such blended acoustic models, thus it is respectfully submitted that Claims 26, and claims 27-30 that depend from it appear to be patentable.

With regard to claim 31, and claims 32-36 that depend from it, like claim 26, these claims all contain a limitation related to blended acoustic models that are weighted in a specified way. Since as is stated above with regard to Claim 26, the Office Action makes no mention of such features, it is respectfully submitted that Claim 31, and claims 32-26 that depend from it, appear to be patentable.

With regard to independent claim 37, and claims 38-39 and 47-52, that depend from it, Claim 37, like claims 26 and 31 just described, recites the use of blended acoustic models that are weighted in a specific way. Since the Office Action makes no mention of such blended models, it is respectfully submitted that claim 37, and all of the claims that depend from it, including claims 38-39 and 47-52, are patentable.

With regard to Paragraph 6's rejection of Claims 53-60, these claims do not appear to be similar in scope with claims 32-36, with the possible exception of claim 59 which has a partial similarity in scope with claim 32. Furthermore, independent claim 53, from which claims 54-60 depend, appears to be distinguished from the combined teachings of Viikki and Baker.

Claim 53 recites a method for training a pronunciation guesser deriving letter-to-phoneme mappings from individual pronunciation-guesser training words where

> "...the using of said letter-to-phoneme mappings includes varying the weight given to a given letter-to-phoneme mapping in the training of the pronunciation guesser as a function of the frequency measure of the word in which such a mapping occurs."

This is clearly distinct from the pronunciation guesser disclosed in Baker. First of all, Baker does not explicitly state how its pronunciation guesser is trained. Second, Baker's pronunciation guesser contains a rules list that associates with each rule mapping a text string to a phoneme output an associated probability reflecting the text strings "frequencies of occurrence in the vocabulary." This is quite different than what is recited in Claim 53 which weights not based on the frequency of occurrence of a character-string-to-phoneme mapping in a vocabulary, but rather on the frequency of occurrence of the word in which a character-string-to-phoneme mapping occurs, when determining the contribution that word's mapping will make in training the pronunciation guesser. For example, if the words being trained are names in a phone book, the frequency of the name in the phone book would be used to weight the word's contribution to the pronunciation guesser's training. This distinction over Baker is important, because it is based on how often a given letter-to-phoneme mapping is used, rather than how often it occurs in a vocabulary. Because of this significant distinction, it is respectfully submitted that Claim 53, and claims 54-60 that depend from it are patentable over the references cited against it.

Paragraph 7 of the Office Action rejected claims 40-43 and 46 over Viikki, Baker, and Sensory. Claims 40-43 and 46 all depend from Claim 37, and, thus, it is respectfully submitted that Claims 40-43 and 46 are patentable for the same reasons, stated above, as is claim 37.

Paragraph 8 of the Office Action stated that claims 11 and 44-45 would be allowable if rewritten in independent form. The undersign thanks the Examiner for this acknowledgement and has responded by rewriting these claims in independent form to speed their allowance.

Regarding Paragraph 9 of the Office Action, from a quick look at the prior art made of record but not relied upon, it does not appear that any of these references would alter the arguments made above.

Two formerly dependent claims have been made independent, thus increasing the number of independent claims by two, requiring an extra independent claim fee of $400.00 under 37 CFR 1.16(h). The is being paid by e-Filing.

You are hereby authorized to bill any additional amounts which may be owing for any matters relating to this response to the Office Action (and any refunds that may be due thereon) to the Porter & Associates deposit account numbered 502110.

Statements appearing above in respect to the disclosures in the cited reference represent the present opinions of the undersigned attorney and, in the event that the Examiner disagrees with any of such opinions, it is respectfully requested that the Examiner specifically indicate those portions of the respective reference providing the basis for a contrary view. If you have any questions or problems with any of the arguments or other contents of this amendment, please contact me at the below phone number or address.

Thus, it is respectfully submitted that with the amendments made above, all of the claims in the above claim listing, that is, Claims 1-60 are allowable and early allowance of them is respectfully requested.

Respectfully Submitted,

Edward W. Porter
Registration No.28,892
Porter & Associates
One Broadway
Suite 600
Cambridge, Mass. 02142
617-494-1722